Existence of an information unit as a postulate of quantum theory

Lluís Masanes[a,1], Markus P. Müller[b], Remigiusz Augusiak[c], and David Pérez-García[d]

[a]H. H. Wills Physics Laboratory, University of Bristol, Bristol BS8 1TL, United Kingdom; [b]Perimeter Institute for Theoretical Physics, Waterloo, ON, Canada N2L 2Y5; [c]Institut de Ciències Fotòniques, Mediterranean Technology Park, 08860 Castelldefels (Barcelona), Spain; and [d]Departamento de Analisis Matemático and Instituto de Matemática Interdisciplinar, Universidad Complutense de Madrid, 28040 Madrid, Spain

Edited by Peter W. Shor, Massachusetts Institute of Technology, Cambridge, MA, and approved August 20, 2013 (received for review March 13, 2013)

Does information play a significant role in the foundations of physics? Information is the abstraction that allows us to refer to the states of systems when we choose to ignore the systems themselves. This is only possible in very particular frameworks, like in classical or quantum theory, or more generally, whenever there exists an information unit such that the state of any system can be reversibly encoded in a sufficient number of such units. In this work, we show how the abstract formalism of quantum theory can be deduced solely from the existence of an information unit with suitable properties, together with two further natural assumptions: the continuity and reversibility of dynamics, and the possibility of characterizing the state of a composite system by local measurements. This constitutes a set of postulates for quantum theory with a simple and direct physical meaning, like the ones of special relativity or thermodynamics, and it articulates a strong connection between physics and information.

postulates of quantum mechanics | physics of information | quantum information

Quantum theory (QT) provides the foundation on top of which most of our physical theories and our understanding of nature sits. This peculiarly important role contrasts with our limited understanding of QT itself, and the lack of consensus among physicists about what this theory is saying about how nature works. Particularly, the standard postulates of QT are expressed in abstract mathematical terms involving Hilbert spaces and operators acting on them, and lack a clear physical meaning. In other physical theories, like special relativity or thermodynamics, the formalism can be derived from postulates having a direct physical meaning, often in terms of the possibility or impossibility of certain tasks. In this work, we show that this is also possible for QT.

The importance of this goal is reflected by the long history of research on alternative axiomatizations of QT, which goes back to Birkhoff and von Neumann (1–3). More recently, initiated by Hardy's work (4), and influenced by the perspective of quantum information theory, there has been a wave of contributions taking a more physical and less mathematical approach (4–8). These reconstructions of QT constitute a big achievement because they are based on postulates having a more physical meaning. However, some of these meanings are not very direct, and a lot of formalism has to be introduced to state them. In this work, we derive finite-dimensional QT from four postulates having a clear and direct physical meaning, which can be stated easily and without the need of heavy formalism. Also, contrary to ref. 5, we write all our assumptions explicitly.

We introduce a postulate named "Existence of an Information Unit," which essentially states that there is only one type of information within the theory. Consequently, any physical process can be simulated with a suitably programmed general purpose simulator. Because the input and output of these simulations are not necessarily classical, this postulate is a stronger version of the Church–Turing–Deutsch Principle (stated in ref. 9). However, it is strictly weaker than the Subspace Axiom, introduced in ref. 4 and used in refs. 5 and 6. An alternative way to read this postulate is that, at some level, the dynamics of any system is substrate independent. Within theories satisfying the Existence of an Information Unit, one can refer to states, dynamics, and measurements abstractly, without specifying the type of system they pertain to; and this is exploited by quantum information scientists, who design algorithms and protocols at an abstract level, without considering whether they will be implemented with light, atoms, or any other type of physical substrate.

More precisely, Existence of an Information Unit states that there is a type of system, the generalized bit or gbit, such that the state of any other system can be reversibly encoded in a sufficient number of gbits (Fig. 1). The reversibility of the encoding implies a correspondence between the states of any system and the states of a multigbit system (or an appropriate subspace). This correspondence also extends to dynamics and measurements: if a given system lacks a particular dynamics, then we can encode its state into a multigbit system, engineer the desired multigbit dynamics, and decode back the resulting state on the given system—effectively implementing the desired dynamics. In classical probability theory, the gbit is the bit, and in QT it is the qubit; but we do not restrict ourselves to these two cases. We postulate that, at some level, everything reduces to information, but we do not specify what information is, except for some requirements that the gbit must satisfy. One of these requirements is "No Simultaneous Encoding," which tells that, if a gbit is used to perfectly encode one classical bit, it cannot simultaneously encode any further information. Two close variants of this are Zeilinger's Principle (10) and Information Causality (11).

Our main contribution is to prove that QT is the only theory satisfying the postulates of Continuous Reversibility, Tomographic

Significance

Despite the enormous success of quantum theory, its significance and meaning are still being debated. In particular, the standard postulates of quantum theory are abstract mathematical statements in terms of complex vectors, self-adjoint operators, etc., and as such they lack a clear physical interpretation. For this reason, it is difficult to assess what they say about the structure of the physical world. In this article, we prove that quantum theory can be formulated through four very simple postulates, each having a direct physical and intuitive meaning. Also, our postulates unveil some connections between physics and information that remain hidden in the standard postulates, thus supporting Wheeler's hypothesis "it from bit."

Author contributions: L.M. designed research; L.M., M.P.M., R.A., and D.P.-G. performed research; and L.M., M.P.M., R.A., and D.P.-G. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: ll.masanes@bristol.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1304884110/-/DCSupplemental.

PHYSICS

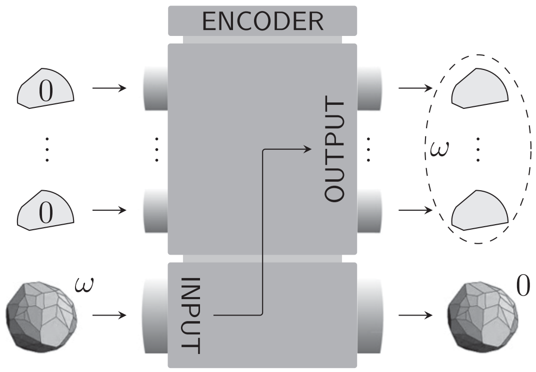

Fig. 1. Encoder. "Coding" is an ideal physical transformation that maps the unknown state ω of an arbitrary system to an n-gbit state in a reversible way and leaves the initial system in a reference state 0. Reversibility means that there is another ideal physical transformation, "decoding," which undoes the above, bringing the arbitrary system back to its original state.

Locality (both introduced in ref. 4), The Existence of an Information Unit, and No Simultaneous Encoding. To prove this, we make use of the classification of state spaces performed in refs. 12 and 13, which shows that quantum state spaces have very special properties. In relation to other work, in ref. 11, it was suggested that Information Causality might be one of the foundational properties of nature. However, our results support that its close variant, No Simultaneous Encoding, might be a better candidate, since it seems to unveil more about the structure of the physical world. Also, our results confirm Zeilinger's idea (10) that the limited amount of information carried by a qubit is a defining property of QT.

A Theory-Independent Formalism

In classical probability theory, no matter how complex a system is, there is a joint probability distribution which simultaneously describes the statistics of all of the measurements that can be performed on a system. In other words, there exists a maximally informative measurement, of which all other measurements are functions. This is not true in QT, and motivated by this, Birkhoff and von Neumann generalized the formalism of classical probability theory to include incompatible measurements (1). This is nowadays called the framework of generalized probability theories (GPTs), or the convex operational framework.

Recently, a lot of interest has been directed to the study of GPTs (4–8, 11, 12, 14–22), with the double aim of reconstructing QT, and exploring what lies beyond. This, in particular, led to the discovery that many features originally thought as specific to QT [such as, for instance, Bell-inequality violation (21), no cloning (15, 22), monogamy of correlations (22), Heisenberg-type uncertainty relations (18, 22), measurement-disturbance trade-offs (15), and the possibility of secret key distribution (23, 24)], are common to most GPTs. In this light, the standard question, "Why does nature seem to be quantum instead of classical?" sounds less appropriate than asking, "Why QT instead of any other GPT?" Here, we answer this question by showing that any GPT different from QT violates at least one of our physically meaningful postulates. In what follows, we derive the formalism of GPTs from the basic notions of state and measurement (a more detailed introduction can be found in *SI Text*).

In QT, states are represented by density matrices. However, how can we represent states in theories that we do not yet know? Let us follow ref. 4. The state of a system is represented by the probabilities of some reference measurement outcomes $x_1, \ldots x_k$ which are called "fiducial":

$$\omega = \begin{bmatrix} p(x_1) \\ \vdots \\ p(x_k) \end{bmatrix} \in \mathcal{S} \subset \mathbb{R}^k. \qquad [1]$$

This list of probabilities has to be minimal but contain sufficient information to predict the probability distribution of all measurements that can be in principle performed on the system. (Note that this is always possible because the list could contain the probabilities corresponding to all measurements. In particular, the list can be infinite, that is, $k=\infty$.) The number of fiducial outcomes k is equal to the dimension of $\mathcal{S}$, as otherwise one fiducial probability would be functionally related to the others, and the list not minimal. We include the possibility that the system is present with certain probability $U \in [0,1]$, which by consistency, is equal to the sum of probabilities for all of the outcomes of a measurement. When the system is absent $(U=0)$, the fiducial outcomes have zero probability; hence the corresponding state (1) is the null vector $\mathbf{0} \in \mathcal{S}$. The subset of normalized states $\mathcal{N} = \{\omega \in \mathcal{S} : U(\omega) = 1\}$ has dimension $k-1$.

By the rules of probability, the set of all of the allowed states $\mathcal{S}$ is convex. Indeed, by preparing the state ω_1 with probability q and ω_2 with probability $1-q$, we effectively prepare the mixed state $q\omega_1 + (1-q)\omega_2$. The "pure states" of $\mathcal{S}$ are the normalized states that cannot be written as mixtures. As an instance, the fiducial outcomes for a qubit can be chosen to be $\sigma_x = 1, \sigma_y = 1, \sigma_z = 1, \sigma_z = -1$, and $U(\omega) = p(\sigma_z = 1) + p(\sigma_z = -1)$. Note that the set of fiducial outcomes need not be unique, nor simultaneously measurable.

In the formalism of GPTs, every convex set can be seen as the state space $\mathcal{S}$ of an imaginary type of system, which, in turn, allows for constructing multipartite states spaces that violate Bell inequalities more (or less) than QT. This illustrates the degree to which this formalism generalizes classical probability theory and QT, and allows us to catch a glimpse on the multitude of alternative theories that we are considering here.

The probability of the measurement outcome x when the system is in the state ω is given by $E_x(\omega)$, where $E_x : \mathbb{R}^k \to \mathbb{R}$ is a linear function satisfying $E_x(\mathcal{S}) \subseteq [0,1]$. To see this, suppose the system is prepared in the mixture $q\omega_1 + (1-q)\omega_2$. Then the relative frequency of an outcome x should not depend on whether the label of the actual preparation ω_k is ignored before or after the measurement. As a result,

$$E_x\big(q\omega_1 + (1-q)\omega_2\big) = qE_x(\omega_1) + (1-q)E_x(\omega_2),$$

which, together with $E_x(\mathbf{0}) = 0$, imply the linearity of E_x.

Physical systems evolve with time. Often, the dynamics of a system can be controlled by adjusting its environment, allowing in this way to engineer different transformations of the system. A transformation can be represented by a map $T : \mathcal{S} \to \mathcal{S}$, which, for the same reason as outcome probabilities E, has to be linear. Sometimes there are pairs of transformations whose composition leaves the system unaffected, independently of its initial state—in this case we say that these transformations are reversible. The set of reversible transformations generated by time-continuous dynamics forms a compact connected Lie group $\mathcal{G}$. Then, the elements of the corresponding Lie algebra are the Hamiltonians of the theory (which in general have nothing to do with Hermitian matrices). Our first postulate imposes that this set of Hamiltonians is sufficiently rich.

The Postulates for QT

Now we are ready to present our axiomatization of QT. The first postulate is motivated by the fact that most physical theories that we know (like, for example, classical mechanics, general relativity, and QT) enjoy time-continuous reversible dynamics.

Postulate 1 (Continuous Reversibility). In any system, for every pair of pure states, one can in principle engineer a time-continuous reversible dynamics that brings one state to the other.

Note that this postulate contains two independent assumptions: reversibility and continuity. As pointed out by Hardy (4), classical probability theory in finite dimensions violates the continuity part of this postulate, because the set of reversible transformations is the group of permutations, which is not connected. Then, if we relax this continuity part, the family of theories satisfying our postulates includes classical probability, but we do not know if it also includes other nonclassical and nonquantum theories.

Now we motivate the second postulate. Let A and B be two systems with fiducial outcomes $x_1,\ldots x_{k_A}$ and $y_1,\ldots y_{k_B}$, respectively. Is there any relation between these and the fiducial outcomes of the composite system AB? The following postulate implies that the set of joint outcomes (x_i,y_j) for all i,j is a fiducial set for the composite system. As a consequence, joint local probabilities (and similarly joint local transformations) can be obtained through the simple tensor-product rule $p(x,y)=(E_x\otimes E_y)(\omega_{AB})$, where

$$\omega_{AB}=\begin{bmatrix} p(x_1,y_1) \\ p(x_1,y_2) \\ \vdots \\ p(x_{k_A},y_{k_B}) \end{bmatrix}\in\mathcal{S}_{AB}\subset\mathbb{R}^{k_A}\otimes\mathbb{R}^{k_B}.$$

This also implies the multiplicativity of dimensions: $k_{AB}=k_Ak_B$.

Postulate 2 (Tomographic Locality). The state of a composite system is completely characterized by the correlations of measurements on the individual components.

The third postulate states the aforementioned existence of the gbit and imposes three properties that it must satisfy.

Postulate 3 (Existence of an Information Unit). There is a type of system (the gbit, with state space denoted $\mathcal{S}_{\text{gbit}}$) such that the state of any system can be reversibly encoded in a sufficiently large number of gbits. Additionally, gbits satisfy the following:

1. State tomography is possible: The state of a gbit can be characterized with a finite number of measurements.
2. All effects are observable: All linear functions $E:\mathcal{S}_{\text{gbit}}\to[0,1]$ correspond to outcomes of measurements that can in principle be performed.
3. Gbits can interact: The group of time-continuous reversible transformations for two gbits contains at least one element that is not product $G_{AB}\neq G_A\otimes G_B$.

Now, let us explain in more detail the content of postulate 3. First, the requirement that the state of any system can be reversibly encoded in a number of gbits is formalized as follows. For any state space $\mathcal{S}$ allowed by the theory, there is a number n, a physical transformation T mapping $\mathcal{S}$ to the state space of n gbits $\mathcal{S}^n_{\text{gbit}}$ (as in Fig. 1), and another physical transformation in the opposite direction $F:\mathcal{S}^n_{\text{gbit}}\to\mathcal{S}$, such that their composition is equal to the identity transformation: $F(T(\omega))=\omega$ for all $\omega\in\mathcal{S}$. This implies that the dimension of $\mathcal{S}^n_{\text{gbit}}$ is not smaller than that of $\mathcal{S}$. If the two dimensions are equal, then the two state spaces are equivalent. However, if the dimension of $\mathcal{S}^n_{\text{gbit}}$ is larger than that of $\mathcal{S}$, then there are states in $\mathcal{S}^n_{\text{gbit}}$ that are not contained in $T(\mathcal{S})$; and for those the transformation F does not work with unit probability. Next, we explain the properties that gbits satisfy.

1. The fact that gbits can be characterized with a finite number of measurements is equivalent to say that the dimension of the state space $\mathcal{S}_{\text{gbit}}$, denoted k_{gbit}, is finite. This may seem contradictory with the fact that, in QT, there is a type of tomography for infinite-dimensional systems. However, these systems have an infinite number of perfectly distinguishable states; hence, after imposing additional constraints (like an upper bound on the energy) the effective Hilbert space is finite, and state tomography becomes possible. However, as a consequence of No Simultaneous Encoding, gbits have only two perfectly distinguishable states.
2. In classical probability theory and QT, all effects correspond to outcomes of measurements. This need not be the case in general, but to single out QT, we have to impose it on gbits. Although in this form this assumption does not have a direct operational meaning, it can be formulated in a way that it does (see ref. 8 or *SI Text*). Unfortunately, this alternative formulation is more cumbersome; hence we avoid it here.
3. Interaction is fundamentally necessary in order not to have an essentially trivial universe. The requirement that any system can be reversibly encoded in gbits implies that, if gbits do not interact among them, then no other system interacts. Postulate 3.3 rules out this possibility.

Postulate 4 (No Simultaneous Encoding). If a gbit is used to perfectly encode one classical bit, it cannot simultaneously encode any further information.

To illustrate postulate 4, let us consider a communication task involving two distant parties, Alice and Bob. Similarly as in the scenario for Information Causality (11), suppose that Alice is given two bits $a,a'\in\{0,1\}$, and Bob is asked to guess one of them. He will base his guess on information sent to him by Alice, encoded in one gbit. Alice encodes the gbit with no knowledge of which of the two bits, a or a', Bob will try to guess. No Simultaneous Encoding imposes that, in a coding/decoding strategy in which Bob can guess a with probability one, he knows nothing about a'. That is, if b, b' are Bob's guesses for a, a' then

$$P\left(b\middle|a,a'\right)=\delta^a_b\Rightarrow P\left(b'\middle|a,a'=0\right)=P\left(b'\middle|a,a'=1\right),$$

where δ^a_b is the Kronecker tensor. A straightforward consequence of this is that $\mathcal{S}_{\text{gbit}}$ contains at most two perfectly distinguishable states. Other consequences are derived below.

Another way to state No Simultaneous Encoding is as follows: suppose that Alice encodes a, a' in the four states $\omega_{a,a'}\in\mathcal{N}_{\text{gbit}}$. If there is an effect E such that $E(\omega_{a,a'})=\delta_{a,0}$, then any effect E' satisfies $E'(\omega_{a,0})=E'(\omega_{a,1})$. As it is illustrated in Fig. 2, this together with All Effects Are Observable (cf. postulate 3.2) imply that all states in the boundary of $\mathcal{N}_{\text{gbit}}$ are pure (first arrow in Fig. 3).

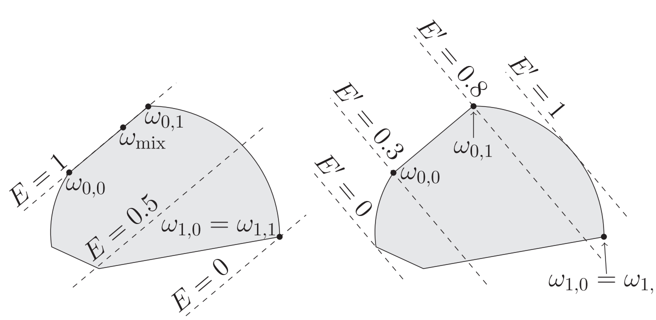

Fig. 2. No simultaneous encoding. This figure shows that there cannot be mixed states in the boundary of $\mathcal{N}_{\text{gbit}}$. If there is one, say ω_{mix}, then this boundary contains a nontrivial face (*Left*). Because all effects are observable, we can decode a with the effect E, which gives probability one for all states inside that facet, and probability zero for some other state(s). By encoding $(a,a')=(0,0),(0,1)$ in two different states inside that face we can perfectly retrieve a through E, while still getting some partial information about a' with another effect E' (*Right*).

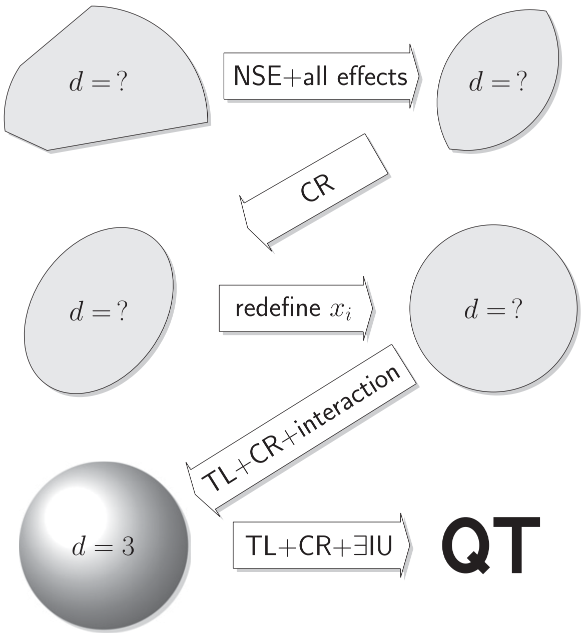

Fig. 3. Summary of the argumentation. This figure synthesizes the proof that the only theory satisfying our four postulates is QT. Each step (represented by an arrow) invokes part of the content of the postulates (specified inside the arrow) and reveals new information about the state space of the generalized bit. Initially (*Top Left*) $\mathcal{N}_{\text{gbit}}$ is an arbitrary convex set with arbitrary dimension $d = k_{\text{gbit}} - 1$, and finally (*Bottom Left*) it is a three-dimensional ball. The first arrow represents the step explained in Fig. 2. The abbreviations CR, TL, ∃IU, NSE, "all effects," and "interaction" refer to Continuous Reversibility, Tomographic Locality and Existence of an Information Unit, No Simultaneous Encoding, All Effects Are Observable, and Gbits Can Interact, respectively.

An interesting remark is that our four postulates, except for part 2 of postulate 3, express the possibility or impossibility of certain tasks. This is very similar in spirit to formulations of the second law of thermodynamics, the principle of equivalence of gravitation and inertia, or the principle of light speed invariance. Contrary, this remains completely hidden in the standard postulates of QT.

Argumentation

Having stated our four postulates, let us now show that the only theory obeying them is QT. In what follows, we present an overview of the proof, whereas its detailed version can be found in *SI Text*. First of all, postulate 3.1 implies that the dimension of the gbit k_{gbit} is finite. Then, Continuous Reversibility associates to any state space $\mathcal{S}$ a group of reversible transformations $\mathcal{G}$, having an invariant scalar product with respect to which all pure states of $\mathcal{S}$ have the same norm. This together with the fact that the boundary of $\mathcal{N}_{\text{gbit}}$ contains only pure states imply that it is an ellipsoid (second arrow in Fig. 3). By setting as the new set of fiducial outcomes the effects corresponding to the principal axes of the ellipsoid (recall that all effects are observable), $\mathcal{N}_{\text{gbit}}$ becomes a Euclidean ball (third arrow in Fig. 3). However, what is the state space of two gbits $\mathcal{S}^2_{\text{gbit}}$? According to Continuous Reversibility, the set of pure states of two gbits can be written as $\{G(\omega\otimes\omega)|G\in\mathcal{G}^2_{\text{gbit}}\}$, where $\mathcal{G}^2_{\text{gbit}}$ is the group of reversible transformations for two gbits, and ω is a pure state of one gbit. The group $\mathcal{G}^2_{\text{gbit}}$ is unknown, but by consistency, it must contain all local transformations,

$$\mathcal{G}_{\text{gbit}}\otimes\mathcal{G}_{\text{gbit}}\subseteq\mathcal{G}^2_{\text{gbit}}, \quad [2]$$

and it must generate states with well-defined probabilities, meaning that

$$(E_x\otimes E_y)(G(\omega\otimes\omega))\in[0,1] \quad [3]$$

holds for all $G\in\mathcal{G}^2_{\text{gbit}}$ and any (local) gbit effects E_x, E_y. The family of all bipartite state spaces satisfying these two consistency requirements was analyzed in ref. 13, and it was shown that, with the exception of the quantum case, all state spaces contain separable states only, and the corresponding groups $\mathcal{G}^2_{\text{gbit}}$ contain product transformations only. However, this is in contradiction with Gbits Can Interact! Hence, the combination of this postulate together with requirements 2 and 3 is very restrictive, and it implies that the Euclidean ball $\mathcal{N}_{\text{gbit}}$ has dimension $k_{\text{gbit}}-1=3$ and $\mathcal{G}_{\text{gbit}}=\mathrm{SO}(3)$ (*SI Text* and ref. 13). This tells us that, locally, gbits are identical to qubits, but it is not clear yet whether multigbit state spaces $\mathcal{S}^n_{\text{gbit}}$ having a nonquantum structure are consistent with our postulates. In ref. 12, all possible joint-state spaces of n systems that are locally qubits are classified, and it is found that the only possibility allowing for nonproduct reversible transformations is multiqubit QT. So gbits must be locally and globally like qubits: $\mathcal{S}^n_{\text{gbit}}$ is the set of 2^n-dimensional density matrices and $\mathcal{G}^n_{\text{gbit}}$ is the adjoint representation of $\mathrm{SU}(2^n)$. Finally, because any state space is reversibly encodable in a multiqubit system, the states, transformations, and measurements of any system can be represented within the formalism of finite-dimensional QT.

Conclusions

Given the controversy around the foundations of QT, it is very natural to seek for modifications and generalizations of QT. In addition, some authors claim that this is necessary to unify the description of quantum and gravitational phenomena (25, 26). Each set of postulates for QT provides a different starting point for this endeavor. For example, starting from the standard postulates, some authors have modified the Schrödinger equation (27), or the field of numbers over which the Hilbert space is defined (28). However, a radically different starting point is provided by our postulates. In *SI Text*, we relax that Gbits Can Interact (postulate 3.3) and characterize the family of theories that emerges (see also ref. 13). It is shown that all these alternative theories, although being not classical, do not contain entanglement and do not violate Bell inequalities. If instead, we relax the continuity part of the Continuous Reversibility Postulate, then the family of theories that emerges includes classical probability theory, but we leave for future research whether other theories are included as well. This seems an important question, because in our construction and others (4), the continuity of the dynamics appears to be the dividing feature between classical probability theory and QT.

A repeated pattern in the history of science is the promotion of a scientific instrument to a model for understanding the world. For instance, there are some proposals for viewing the universe as a giant computer [classical (29) or quantum (30)]. However, what is the physical content of this? Can the dynamics of any system be understood as computation? After all, it is computing its future state. We propose that a requisite for upgrading time evolution to computation is that such time evolution is substrate independent, in the sense that it can be simulated in a system of information units. In this work, we have taken this perspective seriously: we have promoted the Existence of an Information Unit with suitable properties to be a postulate, and we have shown that this together with the very natural postulates of Continuous Reversibility and Tomographic Locality, uniquely determine the full mathematical formalism of QT.

ACKNOWLEDGMENTS. L.M. acknowledges support from CatalunyaCaixa, the European Union (EU) European Research Council Advanced Grant "Nonlocality in Space and Time", the EU "Quantum Interfaces, Sensors and Communication Based on Entanglement" project, the Templeton Foundation, and the FQXi large grant project "Time and the Structure of Quantum Theory." Research at Perimeter Institute for Theoretical Physics is supported in part by the Government of Canada through the Natural Sciences and Engineering Research Council and by the Province of Ontario through the Ministry of Research and Innovation. R.A. acknowledges support from Atomic Quantum Technologies and Spanish Ministerio de Economía y Competitividad through the Juan de la Cierva program. D.P.-G. acknowledges support from Spanish Grants MTM2008-01366 and S2009/ESP-1594.

1. Birkhoff G, von Neumann J (1936) The logic of quantum mechanics. *Ann Math* 37:823–843.
2. Mackey G-W (1963) *The Mathematical Foundations of Quantum Mechanics* (W. A. Benjamin, New York).
3. Alfsen E-M, Shultz F-W (2003) *Geometry of State Spaces of Operator Algebras* (Birkhäuser, Boston).
4. Hardy L (2001) Quantum theory from five reasonable axioms. arXiv:quant-ph/0101012.
5. Dakić B, Brukner Č (2011) Quantum theory and beyond: Is entanglement special? *Deep Beauty: Understanding the Quantum World Through Mathematical Innovation*, ed Halvorson H (Cambridge Univ Press, Cambridge, UK), pp 365-392.
6. Masanes L, Müller M-P (2011) A derivation of quantum theory from physical requirements. *New J Phys* 13:063001.
7. Hardy L (2011) Reformulating and reconstructing quantum theory. arXiv:1104.2066.
8. Chiribella G, D'Ariano G-M, Perinotti P (2011) Informational derivation of quantum theory. *Phys Rev A* 84:012311.
9. Deutsch D (1985) Quantum theory, the Church-Turing principle and the universal quantum computer. *Proc R Soc Lond A Math Phys Sci* 400:97–117.
10. Zeilinger A (1999) A foundational principle for quantum mechanics. *Found Phys* 29:631–643.
11. Pawłowski M, et al. (2009) A new physical principle: Information causality. *Nature* 461:1101–1104.
12. de la Torre G, Masanes L, Short A-J, Müller M-P (2012) Deriving quantum theory from its local structure and reversibility. *Phys Rev Lett* 109(9):090403.
13. Masanes L, Müller M-P, Pérez-García D, Augusiak R (2011) Entanglement and the three-dimensionality of the Bloch ball. arXiv:1111.4060.
14. Chiribella G, D'Ariano G-M, Perinotti P (2010) Probabilistic theories with purification. *Phys Rev A* 81:062348.
15. Barrett J (2007) Information processing in generalized probabilistic theories. *Phys Rev A* 75:032304.
16. Gross D, Müller M, Colbeck R, Dahlsten OC (2010) All reversible dynamics in maximally nonlocal theories are trivial. *Phys Rev Lett* 104(8):080402.
17. Barnum H, et al. (2010) Entropy and information causality in general probabilistic theories. *New J Phys* 12:033024.
18. Oppenheim J, Wehner S (2010) The uncertainty principle determines the nonlocality of quantum mechanics. *Science* 330(6007):1072–1074.
19. Short A-J, Barrett J (2010) Strong nonlocality: A trade-off between states and measurements. *New J Phys* 12:033034.
20. Short A-J, Wehner S (2010) Entropy in general physical theories. *New J Phys* 12:033023.
21. Popescu S, Rohrlich D (1994) Quantum nonlocality as an axiom. *Found Phys* 24(3):379–385.
22. Masanes L, Acín A, Gisin N (2006) General properties of nonsignaling theories. *Phys Rev A* 73:012112.
23. Barrett J, Hardy L, Kent A (2005) No signaling and quantum key distribution. *Phys Rev Lett* 95(1):010503.
24. Masanes L (2009) Universally composable privacy amplification from causality constraints. *Phys Rev Lett* 102(14):140501.
25. Penrose R (1996) On gravitys role in quantum state reduction. *Gen Relativ Gravit* 28:581.
26. Hardy L (2007) Towards quantum gravity: A framework for probabilistic theories with non-fixed causal structure. *J Phys A* 40:3081.
27. Weinberg S (1989) Testing quantum mechanics. *Ann Phys* 194:336–386.
28. Finkelstein D, Jauch J-M, Schiminovich S, Speiser D (1962) Foundations of quaternion quantum mechanics. *J Math Phys* 3(2):207–220.
29. Zuse K (1969) *Rechnender Raum* (Friedrich Vieweg und Sohn, Braunschweig, Germany); (1970) [*Calculating Space*] (Massachusetts Institute of Technology, Project MAC, Cambridge, MA), MIT Technical Translation AZT-70-164-GEMIT. German.
30. Lloyd S (2011) *Programming the Universe* (Random House, New York).

PHYSICS